



OIPE

RAW SEQUENCE LISTING

DATE: 05/14/2002

PATENT APPLICATION: US/09/982,667

TIME: 15:01:09

Input Set : N:\paola\09982667.txt

Output Set: N:\CRF3\05142002\I982667.raw

SEQUENCE LISTING

4 (1) GENERAL INFORMATION:

6 (i) APPLICANT: Prudent, James R.

7 Hall, Jeff G.

8 Lyamichev, Victor I.

10 (ii) TITLE OF INVENTION: Invasive Cleavage Of Nucleic Acids

12 (iii) NUMBER OF SEQUENCES: 69

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Medlen & Carroll, LLP

16 (B) STREET: 220 Montgomery Street, Suite 2200

17 (C) CITY: San Francisco

18 (D) STATE: California

19 (E) COUNTRY: United States Of America

20 (F) ZIP: 94104

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Floppy disk

24 (B) COMPUTER: IBM PC compatible

25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

28 (vi) CURRENT APPLICATION DATA:

C--> 29 (A) APPLICATION NUMBER: US/09/982,667

C--> 30 (B) FILING DATE: 18-Oct-2001

31 (C) CLASSIFICATION:

42 (vii) PRIOR APPLICATION DATA:

34 (A) APPLICATION NUMBER: 08/756,386

35 (B) FILING DATE:

39 (A) APPLICATION NUMBER: US 08/682,853

40 (B) FILING DATE: 12-JUL-1996

43 (A) APPLICATION NUMBER: US 08/599,491

44 (B) FILING DATE: 24-JAN-1996

46 (viii) ATTORNEY/AGENT INFORMATION:

47 (A) NAME: Ingolia, Diane E.

48 (B) REGISTRATION NUMBER: 40,027

49 (C) REFERENCE/DOCKET NUMBER: FORS-02564

51 (ix) TELECOMMUNICATION INFORMATION:

52 (A) TELEPHONE: (415) 705-8410

53 (B) TELEFAX: (415) 397-8338

56 (2) INFORMATION FOR SEQ ID NO: 1:

58 (i) SEQUENCE CHARACTERISTICS:

59 (A) LENGTH: 2506 base pairs

60 (B) TYPE: nucleic acid

61 (C) STRANDEDNESS: double

62 (D) TOPOLOGY: linear

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/982,667

DATE: 05/14/2002

TIME: 15:01:09

Input Set : N:\paola\09982667.txt

Output Set: N:\CRF3\05142002\I982667.raw

```

64      (ii) MOLECULE TYPE: DNA (genomic)
68      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
70 ATGAGGGGGA TGCTGCCCCT CTTTGAGCCC AAGGGCCGGG TCCTCCTGGT GGACGGCCAC      60
72 CACCTGGCCT ACCGCACCTT CCACGCCCTG AAGGGCCTCA CCACCAGCCG GGGGGAGCCG      120
74 GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCCTCAAGG CCCTCAAGGA GGACGGGGAC      180
76 GCGGTGATCG TGGTCTTTGA CGCCAAGGCC CCCTCCTTCC GCCACGAGGC CTACGGGGGG      240
78 TACAAGGCGG GCCGGGCCCC CACGCCGGAG GACTTTCCCC GGCAACTCGC CCTCATCAAG      300
80 GAGCTGGTGG ACCTCCTGGG GCTGGCGCGC CTCGAGGTCC CGGGCTACGA GGCGGACGAC      360
82 GTCCTGGCCA GCCTGGCCAA GAAGGCGGAA AAGGAGGGCT ACGAGGTCCG CATCCTCACC      420
84 GCCGACAAAG ACCTTTACCA GTCCTTTTCC GACCGCATCC ACGTCCTCCA CCCCAGGGGG      480
86 TACCTCATCA CCCC GGCTG GCTTTGGGAA AAGTACGGCC TGAGGCCCCG CAGTGGGGCC      540
88 GACTACCGGG CCCTGACCGG GGACGAGTCC GACAACCTTC CCGGGGTCAA GGGCATCGGG      600
90 GAGAAGACGG CGAGGAAGCT TCTGGAGGAG TGGGGGAGCC TGGAAGCCCT CCTCAAGAAC      660
92 CTGGACCGGC TGAAGCCCGC CATCCGGGAG AAGATCCTGG CCCACATGGA CGATCTGAAG      720
94 CTCCTCTGGG ACCTGGCCAA GGTGCGCACC GACCTGCCCC TGGAGGTGGA CTTGCGCCAA      780
96 AGGCGGGAGC CCGACCGGGA GAGGCTTAGG GCCTTTCTGG AGAGGCTTGA GTTTGGCAGC      840
98 CTCCTCCACG AGTTCGGCCT TCTGGAAGAG CCAAAGCCCC TGGAGGAGGC CCCCTGGCCC      900
100 CCGCCGGAAG GGGCCTTCGT GGGCTTTGTG CTTTCCC GCA AGGAGCCCAT GTGGGCCGAT      960
102 CTTCTGGCCC TGGCCGCGC CAGGGGGGGG CCGGTCCACC GGGCCCCCGA GCCTTATAAA      1020
104 GCCCTCAGGG AGCTGAAGGA GGCGCGGGGG CTTCTCGCCA AAGACCTGAG CGTTCTGGCC      1080
106 CTGAGGGAAG GCCTTGGCCT CCCGCCCGGC GACGACCCCA TGCTCCTCGC CTACCTCCTG      1140
108 GACCCTTCCA ACACCACCCC CGAGGGGGTG GCGCGCGCT ACGGCGGGGA GTGGACGGAG      1200
110 GAGGCGGGGG AGCGGGCCGC CCTTTCCGAG AGGCTCTTCG CCAACCTGTG GGGGAGGCTT      1260
112 GAGGGGGAGG AGAGGCTCCT TTGGCTTTAC CGGGAGGTGG AGAGGCCCTT TTCCGCTGTC      1320
114 CTGGCCCACA TGGAGGCCAC GGGGGTGCGC CTGGACGTGG CCTATCTCAG GGCCTTGTC      1380
116 CTGGAGGTGG CCGAGGAGAT CGCCCGCCTC GAGGCCGAGG TCTTCCGCCT GGCCGGCCAC      1440
118 CCCTTCAACC TCAACTCCCG GGACCAGCTG GAAAGGGTCC TCTTTGACGA GCTAGGGCTT      1500
120 CCCGCCATCG GCAAGACGGA GAAGACCGGC AAGCGCTCCA CCAGCGCCGC CGTCCTGGAG      1560
122 GCCCTCCGCG AGGCCACCCC CATCGTGGAG AAGATCCTGC AGTACCGGGA GCTCACCAAG      1620
124 CTGAAGAGCA CCTACATTGA CCCCTTGCCG GACCTCATCC ACCCCAGGAC GGGCCGCCTC      1680
126 CACACCCGCT TCAACCAGAC GGCCACGGCC ACGGGCAGGC TAAGTAGCTC CGATCCCAAC      1740
128 CTCCAGAAAC TCCCCGTCCG CACCCCGCTT GGGCAGAGGA TCCGCCGGGC CTTTCATCGCC      1800
130 GAGGAGGGGT GGCTATTGGT GGCCCTGGAG TATAGCCAGA TAGAGCTCAG GGTGCTGGCC      1860
132 CACCTCTCCG GCGACGAGAA CCTGATCCCG GTCTTCCAGG AGGGGCGGGA CATCCACACG      1920
134 GAGACCGCCA GCTGGATGTT CGGCGTCCCC CGGGAGGCCG TGGACCCCTT GATGCGCCGG      1980
136 GCGGCCAAGA CCATCAACTT CGGGGTCTCT TACGGCATGT CGGCCACCG CCTCTCCAG      2040
138 GAGCTAGCCA TCCCTTACGA GGAGGCCAG GCCTTCATTG AGCGCTACTT TCAGAGCTTC      2100
140 CCAAGGTGC GGGCCTGGAT TGAGAAGACC CTGGAGGAGG GCAGGAGGCG GGGGTACGTG      2160
142 GAGACCTCT TCGGCCGCCG CCGCTACGTG CCAGACCTAG AGGCCCGGGT GAAGAGCGTG      2220
144 CGGGAGGCGG CCGAGCGCAT GGCTTCAAC ATGCCCGTCC AGGGCACCGC CGCCGACCTC      2280
146 ATGAAGCTGG CTATGGTGAA GCTCTTCCCC AGGCTGGAGG AAATGGGGGC CAGGATGCTC      2340
148 CTTGAGGTCC ACGACGAGCT GGTCTCTGAG GCCCCAAAAG AGAGGGCGGA GGCCGTGGCC      2400
150 CGGCTGGCCA AGGAGGTCAT GGAGGGGGTG TATCCCCTGG CCGTGCCCCT GGAGGTGGAG      2460
152 GTGGGGATAG GGGAGGACTG GCTCTCCGCC AAGGAGTGAT ACCACC      2506
154 (2) INFORMATION FOR SEQ ID NO: 2:
156      (i) SEQUENCE CHARACTERISTICS:
157          (A) LENGTH: 2496 base pairs
158          (B) TYPE: nucleic acid
159          (C) STRANDEDNESS: double

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/982,667

DATE: 05/14/2002

TIME: 15:01:09

Input Set : N:\paola\09982667.txt

Output Set: N:\CRF3\05142002\I982667.raw

```

160          (D) TOPOLOGY: linear
162      (ii) MOLECULE TYPE: DNA (genomic)
166      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
168 ATGGCGATGC TTCCCTCTT TGAGCCCAAA GGCCGCGTGC TCCTGGTGGA CGGCCACCAC      60
170 CTGGCCTACC GCACCTTCTT TGCCCTCAAG GGCTCACCA CCAGCCGCGG CGAACCCGTT      120
172 CAGGCGGTCT ACGGCTTCGC CAAAAGCCTC CTCAAGGCCC TGAAGGAGGA CGGGGACGTG      180
174 GTGGTGGTGG TCTTTGACGC CAAGGCCCCC TCCTTCCGCC ACGAGGCCTA CGAGGCCTAC      240
176 AAGGCGGGCC GGGCCCCCAC CCCGAGGAGC TTTCCCCGGC AGCTGGCCCT CATCAAGGAG      300
178 TTGGTGGACC TCCTAGGCCT TGTGCGGCTG GAGGTTCCCG GCTTTGAGGC GGACGACGTG      360
180 CTGGCCACCC TGGCCAAGCG GCGGAAAAG GAGGGGTACG AGGTGCGCAT CCTCACTGCC      420
182 GACCGCGACC TCTACCAGCT CCTTTCGGAG CGCATCGCCA TCCTCCACCC TGAGGGGTAC      480
184 CTGATCACCC CGGCGTGGCT TTACGAGAAG TACGGCCTGC GCGCGGAGCA GTGGGTGGAC      540
186 TACCGGGCCC TGGCGGGGGA CCCCTCGGAT AACATCCCCG GGGTGAAGGG CATCGGGGAG      600
188 AAGACCGCCC AGAGGCTCAT CCGCGAGTGG GGGAGCCTGG AAAACCTCTT CCAGCACCTG      660
190 GACCAGGTGA AGCCCTCCTT GCGGGAGAAG CTCCAGGCGG GCATGGAGGC CCTGGCCCTT      720
192 TCCCGGAAGC TTTCCAGGT GCACACTGAC CTGCCCCTGG AGGTGGACTT CGGGAGGCGC      780
194 CGCACACCCA ACCTGGAGGG TCTGCGGGCT TTTTGGAGC GGTGGAGTT TGAAGCCTC      840
196 CTCCACAGAG TCGGCCTCCT GGAGGGGCGC AAGGCGGCAG AGGAGGCCCC CTGGCCCCCT      900
198 CCGGAAGGGG CTTTTTGGG CTTTTCCTTT TCCCGTCCCG AGCCCATGTG GGCCGAGCTT      960
200 CTGGCCCTGG CTGGGGCGTG GGAGGGGCGC CTCCATCGGG CACAAGACCC CTTAGGGGGC     1020
202 CTGAGGGACC TTAAGGGGGT GCGGGGAATC CTGGCCAAGG ACCTGGCGGT TTTGGCCCTG     1080
204 CGGGAGGGCC TGGACCTCTT CCCAGAGGAC GACCCCATGC TCCTGGCCTA CTTTCTGGAC     1140
206 CCCTCCAACA CCACCCCTGA GGGGGTGGCC CGGCGTTACG GGGGGGAGTG GACGGAGGAT     1200
208 GCGGGGGAGA GGGCCCTCCT GGCCGAGCGC CTCTTCCAGA CCCTAAAGGA GCGCCTTAAG     1260
210 GGAGAAGAAC GCCTGCTTTG GCTTTACGAG GAGGTGGAGA AGCCGCTTTC CCGGTGTGTTG     1320
212 GCGCGGATGG AGGCCACGGG GGTCCGGCTG GACGTGGCCT ACCTCCAGGC CCTCTCCCTG     1380
214 GAGGTGGAGG CGGAGGTGCG CCAGCTGGAG GAGGAGGTCT TCCGCTGGC CGGCCACCCC     1440
216 TTCAACCTCA ACTCCCGCGA CCAGCTGGAG CGGGTGCTCT TTGACGAGCT GGGCCTGCCT     1500
218 GCCATCGGCA AGACGGAGAA GACGGGGAAA CGCTCCACCA GCGCTGCCGT GCTGGAGGCC     1560
220 CTGCGAGAGG CCCACCCCAT CGTGGACCGC ATCCTGCAGT ACCGGGAGCT CACCAAGCTC     1620
222 AAGAACACCT ACATAGACCC CCTGCCCCGC CTGGTCCACC CCAAGACCGG CCGGCTCCAC     1680
224 ACCCGCTTCA ACCAGACGGC CACCGCCACG GGCAGGCTTT CCAGCTCCGA CCCCACCTG     1740
226 CAGAACATCC CCGTGCGCAC CCCTCTGGGC CAGCGCATCC GCCGAGCCTT CGTGGCCGAG     1800
228 GAGGGCTGGG TGCTGGTGGT CTTGGACTAC AGCCAGATTG AGCTTCGGGT CCTGGCCAC     1860
230 CTCTCCGGGG ACGAGAACCT GATCCGGGTC TTTCAGGAGG GGAGGGACAT CCACACCCAG     1920
232 ACCGCCAGCT GGATGTTCCG CGTTTCCCCC GAAGGGGTAG ACCCTCTGAT GCGCCGGGCG     1980
234 GCCAAGACCA TCAACTTCGG GGTGCTCTAC GGCATGTCCG CCCACCGCCT CTCCGGGGAG     2040
236 CTTTCCATCC CCTACGAGGA GGCGGTGGCC TTCATTGAGC GCTACTTCCA GAGCTACCCC     2100
238 AAGGTGCGGG CCTGGATTGA GGGGACCCTC GAGGAGGGCC GCCGGCGGGG GTATGTGGAG     2160
240 ACCCTCTTCG GCCGCCGGCG CTATGTGCCC GACCTCAACG CCCGGGTGAA GAGCGTGCGC     2220
242 GAGGCGGCGG AGCGCATGGC CTTCAACATG CCGGTCCAGG GCACCGCCGC CGACCTCATG     2280
244 AAGCTGGCCA TGGTGCGGCT TTTCCCCCGG CTTCAGGAAC TGGGGGCGAG GATGCTTTTG     2340
246 CAGGTGCACG ACGAGCTGGT CCTCGAGGCC CCCAAGGACC GGGCGGAGAG GGTAGCCGCT     2400
248 TTGGCCAAGG AGGTCATGGA GGGGGTCTGG CCCCTGCAGG TGCCCTGGA GGTGGAGGTG     2460
250 GGCCTGGGGG AGGACTGGCT CTCCGCCAAG GAGTAG                                2496
252 (2) INFORMATION FOR SEQ ID NO: 3:
254      (i) SEQUENCE CHARACTERISTICS:
255          (A) LENGTH: 2504 base pairs
256          (B) TYPE: nucleic acid

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/982,667

DATE: 05/14/2002

TIME: 15:01:09

Input Set : N:\paola\09982667.txt

Output Set: N:\CRF3\05142002\I982667.raw

```

257         (C) STRANDEDNESS: double
258         (D) TOPOLOGY: linear
260         (ii) MOLECULE TYPE: DNA (genomic)
264         (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
266 ATGGAGGCGA TGCTTCCGCT CTTTGAACCC AAAGGCCGGG TCCTCCTGGT GGACGGCCAC      60
268 CACCTGGCCT ACCGCACCTT CTTGCGCCTG AAGGGCCTCA CCACGAGCCG GGGCGAACCg      120
270 GTGCAGGCGG TCTACGGCTT CGCCAAGAGC CTCCTCAAGG CCCTGAAGGA GGACGGGTAC      180
272 AAGGCCGTCT TCGTGGTCTT TGACGCCAAG GCCCCCTCCT TCCGCCACGA GGCCTACGAG      240
274 GCCTACAAGG CGGGGAGGGC CCCGACCCCC GAGGACTTCC CCCGGCAGCT CGCCCTCATC      300
276 AAGGAGCTGG TGGACCTCCT GGGGTTTACC CGCCTCGAGG TCCCCGGCTA CGAGGCGGAC      360
278 GACGTTCTCG CCACCCTGGC CAAGAAGGCG GAAAAGGAGG GGTACGAGGT GCGCATCCTC      420
280 ACCGCCGACC GCGACCTCTA CCAACTCGTC TCCGACCGCG TCGCCGTCCT CCACCCCGAG      480
282 GGCCACCTCA TCACCCCGGA GTGGCTTTGG GAGAAGTACG GCCTCAGGCC GGAGCAGTGG      540
284 GTGGACTTCC GCGCCCTCGT GGGGGACCCC TCCGACAACC TCCCCGGGGT CAAGGGCATC      600
286 GGGGAGAAGA CCGCCCTCAA GCTCCTCAAG GAGTGGGGAA GCCTGGA AAA CCTCCTCAAG      660
288 AACCTGGACC GGGTAAAGCC AGAAAACGTC CGGGAGAAGA TCAAGGCCCA CCTGGAAGAC      720
290 CTCAGGCTCT CCTTGAGCT CTCCCGGGTG CGCACCAGAC TCCCCCTGGA GGTGGACCTC      780
292 GCCCAGGGGC GGGAGCCCGA CCGGAGGGGG CTTAGGGCCT TCCTGGAGAG GCTGGAGTTC      840
294 GGCAGCCTCC TCCACGAGTT CGGCCTCCTG GAGGCCCCCG CCCCCCTGGA GGAGGCCCCC      900
296 TGGCCCCCGC CGGAAGGGGC CTTCTGTGGC TTCGTCTCTT CCCGCCCGA GCCCATGTGG      960
298 GCGGAGCTTA AAGCCCTGGC CGCCTGCAGG GACGGCCGGG TGCACCGGGC AGCAGACCCC     1020
300 TTGGCGGGGC TAAAGGACCT CAAGGAGGTC CGGGGCCTCC TCGCCAAGGA CCTCGCCGTC     1080
302 TTGGCCTCGA GGGAGGGGCT AGACCTCGTG CCCGGGGACG ACCCATGCT CCTCGCCTAC     1140
304 CTCCTGGACC CCTCCAACAC CACCCCGAG GGGGTGGCGC GCGCTACGG GGGGGAGTGG     1200
306 ACGGAGGACG CCGCCACCG GGCCTCCTC TCGGAGAGGC TCCATCGGAA CCTCCTTAAG     1260
308 CGCCTCGAGG GGGAGGAGAA GCTCCTTTGG CTCTACCACG AGGTGGA AAA GCCCCTCTCC     1320
310 CGGGTCTCTG CCCACATGGA GGCCACCGGG GTACGGCTGG ACGTGGCCTA CCTTCAGGCC     1380
312 CTTTCCCTGG AGCTTGCGGA GGAGATCCGC CGCCTCGAGG AGGAGGTCTT CCGCTTGGCG     1440
314 GGCCACCCCT TCAACCTCAA CTCCCGGGAC CAGCTGGA AA GGGTGCTCTT TGACGAGCTT     1500
316 AGGCTTCCCG CCTTGGGGAA GACGCAAAAG ACAGGCAAGC GCTCCACCAG CGCCGCGGTG     1560
318 CTGGAGGCCC TACGGGAGGC CCACCCATC GTGGAGAAGA TCCTCCAGCA CCGGGAGCTC     1620
320 ACCAAGCTCA AGAACACCTA CGTGGAACCC CTCCCAAGCC TCGTCCACCC GAGGACGGGC     1680
322 CGCCTCCACA CCCGCTTCAA CCAGACGGCC ACGGCCACGG GGAGGCTTAG TAGCTCCGAC     1740
324 CCCAACCTGC AGAACATCCC CGTCCGACCC CCTTGGGGCC AGAGGATCCG CCGGGCCTTC     1800
326 GTGGCCGAGG CGGGTTGGGC GTTGGTGGCC CTGGACTATA GCCAGATAGA GCTCCGCGTC     1860
328 CTCGCCACC TCTCCGGGGA CGAAAACCTG ATCAGGGTCT TCCAGGAGGG GAAGGACATC     1920
330 CACACCCAGA CCGCAAGCTG GATGTTCTGG GTCCCCCGG AGGCCGTGGA CCCCCTGATG     1980
332 CGCCGGGCGG CCAAGACGGT GAACTTCGGC GTCTCTACG GCATGTCCGC CCATAGGCTC     2040
334 TCCCAGGAGC TTGCCATCCC CTACGAGGAG GCGGTGGCCT TTATAGAGGC TACTTCCAAA     2100
336 GCTTCCCCAA GGTGCGGGCC TGGATAGAAA AGACCCTGGA GGAGGGGAGG AAGCGGGGCT     2160
338 ACGTGGAAC CCTCTTCGGA AGAAGGCGCT ACGTGCCCGA CCTCAACGCC CGGGTGAAGA     2220
340 GCGTCAGGGA GGCCGCGGAG CGCATGGCCT TCAACATGCC CGTCCAGGGC ACCGCCGCCG     2280
342 ACCTCATGAA GCTCGCCATG GTGAAGCTCT TCCCCGCTT CCGGGAGATG GGGGCCGCA     2340
344 TGCTCCTCCA GGTCCACGAC GAGCTCCTCC TGGAGGCCCC CCAAGCGCGG GCCGAGGAGG     2400
346 TGGCGGCTTT GGCCAAGGAG GCCATGGAGA AGGCCTATCC CCTCGCCGTG CCCCTGGAGG     2460
348 TGGAGGTGGG GATGGGGGAG GACTGGCTTT CCGCAAGGG TTAG      2504
350 (2) INFORMATION FOR SEQ ID NO: 4:
352         (i) SEQUENCE CHARACTERISTICS:
353         (A) LENGTH: 832 amino acids

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/982,667

DATE: 05/14/2002

TIME: 15:01:09

Input Set : N:\paola\09982667.txt

Output Set: N:\CRF3\05142002\I982667.raw

```

354      (B) TYPE: amino acid
355      (C) STRANDEDNESS: single
356      (D) TOPOLOGY: linear
358      (ii) MOLECULE TYPE: protein
362      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
364      Met Arg Gly Met Leu Pro Leu Phe Glu Pro Lys Gly Arg Val Leu Leu
365      1          5          10          15
367      Val Asp Gly His His Leu Ala Tyr Arg Thr Phe His Ala Leu Lys Gly
368      20          25          30
370      Leu Thr Thr Ser Arg Gly Glu Pro Val Gln Ala Val Tyr Gly Phe Ala
371      35          40          45
373      Lys Ser Leu Leu Lys Ala Leu Lys Glu Asp Gly Asp Ala Val Ile Val
374      50          55          60
376      Val Phe Asp Ala Lys Ala Pro Ser Phe Arg His Glu Ala Tyr Gly Gly
377      65          70          75          80
379      Tyr Lys Ala Gly Arg Ala Pro Thr Pro Glu Asp Phe Pro Arg Gln Leu
380      85          90          95
382      Ala Leu Ile Lys Glu Leu Val Asp Leu Leu Gly Leu Ala Arg Leu Glu
383      100         105         110
385      Val Pro Gly Tyr Glu Ala Asp Asp Val Leu Ala Ser Leu Ala Lys Lys
386      115         120         125
388      Ala Glu Lys Glu Gly Tyr Glu Val Arg Ile Leu Thr Ala Asp Lys Asp
389      130         135         140
391      Leu Tyr Gln Leu Leu Ser Asp Arg Ile His Val Leu His Pro Glu Gly
392      145         150         155         160
394      Tyr Leu Ile Thr Pro Ala Trp Leu Trp Glu Lys Tyr Gly Leu Arg Pro
395      165         170         175
397      Asp Gln Trp Ala Asp Tyr Arg Ala Leu Thr Gly Asp Glu Ser Asp Asn
398      180         185         190
400      Leu Pro Gly Val Lys Gly Ile Gly Glu Lys Thr Ala Arg Lys Leu Leu
401      195         200         205
403      Glu Glu Trp Gly Ser Leu Glu Ala Leu Leu Lys Asn Leu Asp Arg Leu
404      210         215         220
406      Lys Pro Ala Ile Arg Glu Lys Ile Leu Ala His Met Asp Asp Leu Lys
407      225         230         235         240
409      Leu Ser Trp Asp Leu Ala Lys Val Arg Thr Asp Leu Pro Leu Glu Val
410      245         250         255
412      Asp Phe Ala Lys Arg Arg Glu Pro Asp Arg Glu Arg Leu Arg Ala Phe
413      260         265         270
415      Leu Glu Arg Leu Glu Phe Gly Ser Leu Leu His Glu Phe Gly Leu Leu
416      275         280         285
418      Glu Ser Pro Lys Ala Leu Glu Glu Ala Pro Trp Pro Pro Pro Glu Gly
419      290         295         300
421      Ala Phe Val Gly Phe Val Leu Ser Arg Lys Glu Pro Met Trp Ala Asp
422      305         310         315         320
424      Leu Leu Ala Leu Ala Ala Ala Arg Gly Gly Arg Val His Arg Ala Pro
425      325         330         335
427      Glu Pro Tyr Lys Ala Leu Arg Asp Leu Lys Glu Ala Arg Gly Leu Leu
428      340         345         350

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/982,667

DATE: 05/14/2002
TIME: 15:01:10

Input Set : N:\paola\09982667.txt
Output Set: N:\CRF3\05142002\I982667.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; N Pos. 4,5,181,182,190,366,617,628,685,714,722,738,784,1022,1029
Seq#:7; N Pos. 1038,1053,1098,1105,1206,1227,1244,1251,1252,1253,1350,1380
Seq#:7; N Pos. 1497,1530,1569,1572,1641,1653,1655,1770,1812,2319,2346,2396
Seq#:8; Xaa Pos.2,63,109,186,205,209,227,228,233,240,243,244,247,260,290
Seq#:8; Xaa Pos.329,336,340,368,414,417,418,431,551,605,773,794,798,823,833
Seq#:57; N Pos. 18

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/982,667

DATE: 05/14/2002

TIME: 15:01:10

Input Set : N:\paola\09982667.txt

Output Set: N:\CRF3\05142002\I982667.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:976 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:0
L:985 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:48
L:994 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:96
L:1009 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:176
L:1012 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:192
L:1015 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:208
L:1018 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:224
L:1021 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:240
L:1024 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:256
L:1030 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:288
L:1036 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:320
L:1039 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:336
L:1042 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:352
L:1051 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:400
L:1054 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:416
L:1078 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:544
L:1087 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:592
L:1120 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:768
L:1123 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:784
L:1129 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:816
L:1132 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8 after pos.:832
L:2271 M:246 W: Invalid value of Alpha Sequence Header Field, [TOPOLOGY:], SeqNo=56